

Mustafa Sezer Başar

Business & Data Analysis

Ankara, Türkiye · +90 533 024 11 76 · mustafasezerbasar@gmail.com · github.com/mustafa-sezer-basar · linkedin.com/in/mustafa-sezer-basar

PROFESSIONAL SUMMARY

Business Administration graduate. Ran the 15-person business he built from scratch in Cappadocia and wrote its software himself; today he leads a three-person technology team at a retail group. Six of his projects are open source.

HIGHLIGHTS

95% reduction in operational errors · 200% ad ROI (profit over spend) · 2 AI products built end to end · 15-person team built from scratch · SEO: #72 → #2 in local search (keyword-pool avg.)

EXPERIENCE

Paulmark — Database Analyst — Business Development & Strategy

June 2026 – Present · Ankara

- Leads a three-person technology team, moving a Nebim V3 and Power BI retail stack off Excel-based processes.
- Automated the monthly employee-commission calculation: sales data loaded from Nebim V3 or by manual entry runs through a single computation flow. Also added a machine learning model that produces suggestions from past sales.
- Compared the company's sales against the TÜİK retail sales index and built a sales forecast model driven by the index's three-month projection.
- Closed the gap between the teams' own Excel files and the master data; inconsistencies used to surface only once they had grown.
- Automated the Excel-based data reconciliation of the Kazakhstan operation; scheduling it is in progress.
- Builds department-level solutions on POS and accounting data; in a family company, agreeing on definitions comes before the technical work.
- Researching solutions to the cold start problem in the store-product distribution application.

Myrco Spa Center (Barceló Cappadocia) — Co-Founder / Operations Lead

September 2024 – February 2025 · Cappadocia

- Ran the spa centre inside the Barceló Cappadocia hotel as a subcontractor; built the 15-person team from scratch.
- Managed cash flow, break-even analysis and financial projections to keep the business sustainable.
- Cut team-caused errors 95% in two months with a Spa Management app he built; error = an unauthorised entry counted on camera, or a session off the sold duration by ±5%.

Cappadocia High Travel Agency — Sales & Marketing Lead

April 2023 – July 2024 · Cappadocia

- Moved the average Google local search rank across a Cappadocia keyword pool ('travel', 'travel agency' and similar) from 72 to 2 in three months through data-driven SEO.
- Ran the allocation of the marketing budget across channels, the break-even calculation and cash-flow tracking.
- Planned and ran the digital advertising campaigns; the return on advertising investment came close to 200%.
- Planned regional tours and ran field operations for international events such as "Echoes from Agarthā".

Digital Consulting (Freelance) — Business & Technology Consultant

2021 – Present · SMB clients

- Has served 13 SMEs across sectors since 2021 with digital advertising management (SEM, social) and website development/optimisation; engagements ranged from one-off improvements to ongoing ad management.

PROJECTS

- **Tax Audit System with Rule Extraction from Legislation** — The rule engine of a field audit system built for an audit firm: an end-to-end working component that reads legislation with RAG and turns it into audit rules. The corpus is self-sourced: a collector he built pulls statutes and high-court rulings into a database, and the RAG layer works on top of it. Rules are held as JSON with their statute reference and validity period, pass a human review gate, and are then applied deterministically by a Go service. The LLM is not in the hot path — it runs offline — so every audit run stays reproducible and auditable. Taxpayer data never leaves the system: the external language model only ever sees public legislation. (*Python (LLM), Go, Next.js, MySQL, Redis*)
- **Multi-Country Retail Data Platform — Target Architecture** — Target data architecture and staged migration path for a multi-country retail group: six documents, five decision records. The diagnosis is that volume is not the problem — ~150 million rows a year is a size one partitioned relational database carries comfortably. The problems are definition (one word, different numbers), authority (no single publisher) and latency (inconsistency is noticed only once it is large). So most of the architecture is a list of things deliberately not built: distributed processing, a data lake and streaming infrastructure are all out of scope with reasons. Power BI, Excel and language-model access all bind to the same semantic layer — a definition that does not live in one place lives in three and produces three different numbers. The model gets no free SQL; it gets metric-shaped tools and runs under the identity of whoever asked. The binding constraint on the migration order is adoption: no old path is closed before the new one is seen working. Design stage — no migration phase is complete, and the volume figures are derived from the shape of the business, not measured. (*Mimari tasarım, katmanlı ilişkisel veri ambarı, semantik katman, ADR*)
- **KVKK Compliance Engine** — Full-stack system encrypting employee PII field by field with AES-256-GCM: a key per subject makes erasure the destruction of that key (crypto-shredding); plus role-aware masking, a hash-chained audit log and 64 adversarial tests generated from the policy table. Compliance claims come from an evidence report that fails the build when dirty. (*TypeScript, Express, node:sqlite, Node crypto (AES)*)
- **Competitor Rate Monitoring & Dynamic Pricing System** — End-to-end system written to collect rates daily from several booking channels (Booking, Hotels and others), keep them as a time series, and turn them into a price recommendation: collection runs through a Chrome extension in a real browser session after the headless approach was stopped by bot protection, with history in PostgreSQL behind an Express API and a React + Recharts interface. The recommendation's distinguishing input is competitor occupancy — rates are public, occupancy is not. It binary-searches the channel's availability query by raising the room count, which reveals the remaining inventory the competitor is willing to show, then divides by their published room count. No reservation is ever created; it only reads the availability response any customer can see. To find out how accurate the method actually is, wrote an open-source study: in a simulated market where true occupancy is known, the naive reading is 35 points out and a censored-regression estimator 12 — but when the revenue effect was measured, the gain did not beat a one-line control needing no collector at all, and it was reported that way. (*TypeScript, Express, node:sqlite, Chrome Extension*)
- **Schedule Risk and Critical Path Analysis** — A simulation study measuring how likely a project schedule is to meet its committed date. On a twenty-activity ERP migration schedule the deterministic critical path gives 105.2 days; simulated 40,000 times from the same estimates, that date is met in only 34% of runs. The gap comes from the shape of the network rather than the estimates: every merge point waits for the slowest of several parallel paths. Label-free control networks show the project finishing 12.6 days later purely from adding parallel paths, with no task lengthened. The criticality index was also measured to tie six activities at 1.0 while their influence on the finish differed 4.6-fold. Interventions were compared with common random numbers; the mechanism in the pre-registered hypothesis was refuted and reported as such. (*Python (NumPy, SciPy, Monte Carlo)*)
- **Data Quality Audit & Reporting Tool** — Tool that cleans messy tabular data into a multi-sheet Excel report — and, more to the point, measures the cleaner's own error rate. It generates a synthetic file with known defects and scores against ground truth, which makes two outcomes visible that cleaning tools never report: errors missed in silence, and correct cells destroyed. Against an ordinary twenty-line pandas cleaner: 10 silent errors versus 361. The measurement found three defects in the tool itself (61 correct salaries multiplied by twelve, 28 required fields silently blank, 8 malformed emails with no log entry), none of them visible in the output. A pay-gap analysis reads 3.37% on dirty data and 3.86% after cleaning, against a true 3.85%. 134 tests. (*Python, pandas, openpyxl*)
- **Face-Recognition Self Check-in System (Portal Cappadocia)** — Commercially built product that lets a guest check in without going to the front desk. It reached production readiness but was never deployed: the hotel management system used in the region was closed to outside integration. Three steps: the guest uploads an ID or passport and its text is read with Tesseract, takes a photo of themselves, and the two faces are matched with face_recognition — then location is verified, because a face match proves who someone is, not that they are at the hotel. Face data is a special category under Turkish data protection law, so the flow opens with an explicit consent gate and a guest who declines checks in face to face instead: consent without an alternative is not consent. No photograph is kept: once the comparison is done, only the numeric representation of the face remains, and the hotel sets how long it is held — under the Turkish data protection act the hotel is the data controller, so that decision belongs there. There is no liveness check; the system can be fooled by a photograph held up to the camera, and the location check does not close that gap. (*Python, Tesseract OCR, face_recognition*)

SKILLS

- **Data Analytics & Reporting:** Data quality and reconciliation checks, Metric definition and shared dictionary (consistency across reports), Source-to-report data flow (ingestion, transformation, publication), Nebim V3 (retail ERP) data model, Measurement design: control group, placebo, pre-registered hypothesis
- **Data Science & AI:** Python (pandas, scikit-learn, NumPy), RAG architecture, LLM API integration, structured output, Multi-Criteria Decision Making (WASPAS, VIKOR, MACBETH), Power BI, advanced Excel (reporting and data preparation), Computer Vision (face_recognition, Tesseract OCR)
- **Web & Full-Stack:** JavaScript / TypeScript, React, Next.js, Express.js, REST API, Puppeteer
- **Finance & Strategy:** Financial modeling, forecasting, ROI/NPV/IRR analysis, Break-even analysis, risk management, Portfolio optimization (Markowitz), Monte Carlo simulation
- **Databases:** SQL (joins, window functions, indexing), PostgreSQL (pgvector), MySQL, SQLite, Prisma (ORM), Redis (caching)
- **Infrastructure & Operations:** End-to-end bare-metal setup: reverse proxy, TLS certificate management, service management with pm2, DDoS protection and rate limiting via Cloudflare, PostgreSQL provisioning, logging and alerting, Self-hosting (Node + MySQL, cPanel/Hostinger), Git, GitHub
- **Security & Compliance:** AES encryption, RBAC, PII masking, Audit logging, KVKK/GDPR compliance

EDUCATION

Hacettepe University — BA in Business Administration

August 2022 – July 2026 · GPA 3.12/4.00 · English-medium · Graduated

Middle East Technical University (METU) — Civil Engineering

September 2018 – August 2022 · English-medium · Transferred to Business Administration at Hacettepe University

LEADERSHIP & COMMUNITIES

- Founding team member, Hacettepe Management Consulting Club (HMCC)
- Debate competitor at METU; special speaker award at the novice tournament
- Member of the METU Fine Arts Society a cappella choir

LANGUAGES

Turkish (Native) · English (Advanced — English-medium undergraduate education)